RTB HEADER BIDDER EVIDENCE
–
EXPLANATORY DOCUMENT

I. Background

1. Brave has conducted this research to understand the personal data flows in Google's Authorized Buyers RTB system.

2. The Data Subject is Dr Johnny Ryan, the complainant before the Data Protection Commissioner. Dr Ryan is Chief Policy & Industry Relations Officer at Brave. We shall refer to Dr Ryan as the Data Subject herein. This document provides an analysis of automatically recorded log of the Data Subject's web browser's activity over the course of approximately one hour of web browsing.

3. This log includes all items (including web pages and their component parts, files, etc.) that the Data Subject's web browser was instructed to load by the web sites that he visited. We refer to it as the "network log" hereafter.

4. Our analysis of the network log shows that the Data Subject's personal data has been processed in Google's Authorized Buyers RTB system. It further shows that Google has also facilitated the sharing of personal data about the Data Subject between other companies.

5. The number of instances in which Google identifiers about the Data Subject are captured in the Data Subject's browser network log, both by Google and by others,

reveals the speed and scale of server-to-server RTB broadcasts. As is explained below, this is a fraction of the totality of personal data broadcast in the RTB server-to-server broadcasts during the Data Subject's brief browsing session.

II. Methodology

6. The Data Subject was using Google's Chrome browser, with no extensions installed. The Data Subject had no logins, cookies, or browsing history on the device. The user in affect appeared as a new user without any background at the start of the browsing session.

III. Proof that the Data Subject's personal data were broadcast by Google in Authorized Buyers bid requests

7. The network log reveals that the data subject's personal data was broadcast in bid requests. In particular, Google's encrypted user identifier "google_gid" about the Data Subject were present in the network log.

8. These identifiers are not anonymous. Rather they are pseudonymous.[1] As Recital 30 of the GDPR summarises:

> "Natural persons may be associated with online identifiers provided by their devices, applications, tools and protocols, such as internet protocol addresses, cookie identifiers or other identifiers such as radio frequency identification tags. This may leave traces which, in particular when combined with unique identifiers and other information received by the servers, may be used to create profiles of the natural persons and identify them."[2]

[1] See GDPR 4 (5), and Recital 26.
[2] GDPR, Recital 30.

9. The network log only records network traffic that passes through the Data Subject's computer. It does not therefore show the initial server-to-server RTB auction from Google. However, it does show the consequences, in particular the proliferation of Google's user identifiers for the Data Subject among companies that are not Google. In fact, the network log shows that Google identifiers for the Data Subject are first used by companies other than Google, rather than by Google itself. The various companies using Google's identifiers therefore must have received them in RTB (server to server) post-bid request cookie matching. We have provided an accompanying sequence diagram to show the stages of this process.

10. In our analysis of only one hour of web browsing by the Data Subject, "google_gid" identifiers about the Data Subject were used 318 times in network traffic from ten companies participating in RTB auctions. These ten companies each obtained a "google_gid" identifier from a "cookie matching" process. Cookie matching allows companies to cross-reference their identifiers for the same person. This in turn allows participants who have previously profiled the Data Subject to know that the Data Subject is the same person that they have previously profiled so that they can update their profile about him and decide whether to bid for his attention.[3]

11. The ten companies were only allowed to use Cookie Matching to receive a "google_gid" from Google because they had each won RTB auctions.[4] We cannot know how many other companies participated in these auctions, and thereby received the Data Subject's personal data in bid requests from Google.[5] However, Google's documentation attests that at least 833 or 2,033 companies receive personal data from it during RTB auctions.[6] The Data Subject is therefore likely to

[3] See "Cookie matching", Authorized Buyers, 25 June 2019 (URL: https://developers.google.com/authorized-buyers/rtb/cookie-guide).
[4] See "Cookie matching", Authorized Buyers, 25 June 2019 (URL: https://developers.google.com/authorized-buyers/rtb/cookie-guide).
[5] For detail on the personal data contained in bid requests see "Authorized Buyers Real-Time Bidding Proto", Google (URL: https://developers.google.com/authorized-buyers/rtb/realtime-bidding-guide).
[6] "Ad Manager and Ad Exchange program policies: Ad Technology Providers", Google Ad Manager Help (URL: https://support.google.com/admanager/answer/9012903) for the list of 833 companies that

have had his personal data profiled by a vast array of companies, rather than just the ten recorded in the network log.

12. The network log does not show Google's server-to-server "cookie matching" with other companies, because the data subject's computer is not privy to what happens between servers. However, we are able to understand what does occur between servers because of Google's cookie matching documentation, which sets out in detail how this system sends additional personal data.[7] The further identifiers and other personal data broadcast from Google's ad exchange server to auction participants' servers are set out in detail in Google's "Authorized Buyers" protocol.[8] (This protocol was presented in evidence with the initial complaint.)

IV. Google's new real-time bidding GDPR "workaround"

13. The network log also shows that Google allowed companies to work around GDPR protections that Google purports to observe. Google's documentation says (emphasis added):

> In another component of Google's cookie matching code, called pixel matching, Google algorithmically selects **an additional buyer** whose cookie can be matched with the Google User ID. Google then places a match tag onto the impression, and includes the chosen buyer's URL in the

---

Google shares personal data with in the European Economic Area. See "Ad Exchange Certified External Vendors", Google Developers (https://developers.google.com/third-party-ads/adx-vendors) for the list of 2,033 companies that Google shared personal data with – it is not clear whether this list also applies in the European Economic Area.

[7] "Ad Manager and Ad Exchange program policies: Ad Technology Providers", Google Ad Manager Help (URL: https://support.google.com/admanager/answer/9012903) for the list of 833 companies that Google shares personal data with in the European Economic Area. See "Ad Exchange Certified External Vendors", Google Developers (https://developers.google.com/third-party-ads/adx-vendors) for the list of 2,033 companies that Google shared personal data with – it is not clear whether this list also applies in the European Economic Area.

[8] Authorized Buyers Real-Time Bidding Proto, Authorized Buyers, last updated 10 July 2019 (URL: https://developers.google.com/authorized-buyers/rtb/realtime-bidding-guide).

> match tag. ...Google places the match tag on the page, which combines a buyer-supplied URL with the Google User ID (the google_gid parameter) and a new google_push parameter.[9]

Nevertheless, the network log reveals that the reference to a singular "additional buyer" is false. In fact, Google allowed not only one additional party, but many, to match with the Google User ID about the Data Subject.

14. In the same document, Google claims that "Google prohibits multiple buyers from joining their match tables."[10] In fact, the network log further reveals that Google allowed multiple parties to match with their identifiers for the data subject with each other.

15. The network log shows that Google creates "cookie_push.html" web pages (Google Push Pages hereafter). Google Push Pages display no visible content. These pages have URLs (page names) that are unique to the Data Subject. This in turn allows companies to pseudonymously identify the Data Subject, in circumstances where this would not otherwise be possible. Thus, the network traffic triggered by these Google Push Pages provides a hidden mechanism for the sharing of the Data Subject's personal data between Google and RTB companies.[11]

16. Analysis of these Google Push Page pages about the Data Subject reveals that RTB companies sent requests to Google that contained not only the "google_gid"

---

[9] "Cookie matching", Authorized Buyers, 25 June 2019 (URL: https://developers.google.com/authorized-buyers/rtb/cookie-guide).

[10] "Cookie matching", Authorized Buyers, 25 June 2019 (URL: https://developers.google.com/authorized-buyers/rtb/cookie-guide).

[11] Google has a default: "If a publisher doesn't engage with these controls to choose their own list, we will apply a list of commonly used Ad Technology Providers." What the publisher is unlikely to realise is that there are 199 companies on this list. See "Ad Manager and Ad Exchange program policies", Google Ad Manager Help (URL: https://support.google.com/admanager/answer/9012903?hl=en).

identifier about the Data Subject, but also a new identifier called “google_push”. Other data were also sent.[12]

17. Like the “google_gid” identifier, the “google_push” identifier was previously given to these companies by Google to identify the Data Subject. However, there is an important difference between the two identifiers. The “google_gid” identifier is specific to each company that receives it. This means that these companies are less likely to be able to cross-reference what they learn about the Data Subject from Google with each other. The “google_push” identifier, however, is common to multiple companies who receive it. The Push Pages provided herewith show that several companies received the same google_push identifier about the Data Subject from Google. This allows them to cross-reference their profiles about the Data Subject.

18. This is contrary to Google’s prohibition against “multiple buyers from joining their match tables”, and declaration that “bidders are expected to support the above principles, and to safeguard user privacy in their implementations”.[13] It also appears to be a workaround to the data-trading limits, which Google believes apply under the GDPR. Google introduced new limitations on its RTB partners’ access to identifiers on the day that the GDPR was applied, which the hidden Push Pages circumvent.[14]

---

[12] Other company specific fields include google_nid (Network ID), google_ula (an optional timestamp, and userlistid) and the option to append custom parameters to the redirect URIs like “p1=v1&p2=v2”, which allow nearly unlimited possibilities for segment identifiers. See "Cookie matching", Authorized Buyers, Google (URL: https://developers.google.com/authorized-buyers/rtb/cookie-guide).

[13] “Cookie matching”, Authorized Buyers, 25 June 2019 (URL: https://developers.google.com/authorized-buyers/rtb/cookie-guide).

[14] "Google Sharply Limits DoubleClick ID Use, Citing GDPR", AdExchanger, 27 April 2018 (URL: https://adexchanger.com/platforms/google-sharply-limits-doubleclick-id-use-citing-gdpr/). Google made a public document in September 2018 that “We no longer populate encrypted UserID and PartnerID fields in Data Transfer for events associated with EEA users recorded in Campaign Manager and Display & Video 360” and that “We removed encrypted cookie IDs and list names (if used) from the Data Transfer file for all global bid requests to Authorized buyers”.
See “Important changes to data transfer”, Google Campaign Manager Help, 5 September 2018 (URL: https://support.google.com/dcm/answer/9006418?hl=en).

19. In response to these requests, Google triggers its own requests to these companies as the Push Page loads. It sends a “google_hm” identifier and other fields about the Data Subject. This mechanism allows all parties involved in the loading of a Push Page to tie their identifiers about the Data Subject together.[15]

20. The network log reveals that over the course of only one hour of web browsing by the Data Subject, Google created at least six unique Push Page pages and eleven duplicate Push Pages that triggered unique data transfers about the Data Subject.[16] Eight companies other than Google were active on one or more of these pages. Google_push identifiers for this specific Data Subject were used 278 times in network traffic. Copies of these pages are provided as evidence. Other parties, such as vendors of browser extensions, are able to access Google Push Pages, further leaking data about the Data Subject.

21. Push Pages undermine Google’s purported data protection measures. They are also vulnerable to abuse by other parties. We are aware that companies other than Google have used the Push Page mechanism to establish their own Push Pages to share data with their own business partners. This appears to happen without Google’s knowledge. The loss of control over personal data in Google’s RTB system is again evident, and it is clear that Google’s policies have offered no protection.

Dr Johnny Ryan
Brave, Inc.
02 September 2019

[15] “Hosted Match Tables” Authorized Buyers, Last updated June 25, 2019 (URL: https://developers.google.com/authorized-buyers/rtb/cookie-guide#benefits-of-hosted-match-tables)

[16] Sixteen Google Push Pages were created, of which nine were unique. See copies of these pages submitted in evidence.